

(2)

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-032422

(43)Date of publication of application : 28.01.2000

(51)Int.Cl.

H04N 7/15
G06F 3/16

(21)Application number : 10-193145

(71)Applicant : NEC CORP

(22)Date of filing : 08.07.1998

(72)Inventor : KAMURA YUKARI

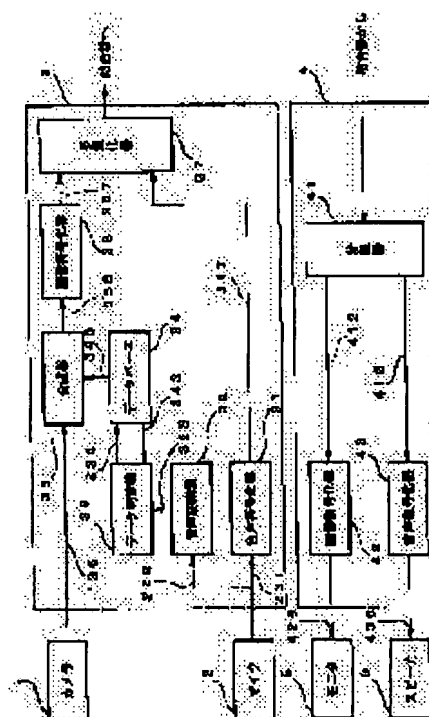
(54) IMAGE TRANSMITTER

(57)Abstract:

PROBLEM TO BE SOLVED: To provide an image transmitter that recognizes voice data to be transmitted selects data corresponding to the voice data from a database, and transmits synthesized data between transmission video data and the voice data.

SOLUTION: A voice coder 31 codes a voice signal from a microphone 2, the coded signal is fed to a multiplexer 37, and a voice recognition device 32 recognizes the signal and gives the signal to a data discrimination device 33. The data discrimination device 33 selects the data corresponding to the voice signal from a database 34 and gives the selected data to a synthesizer 35. The synthesizer 35 synthesizes data with a video image from a camera 1 an image coder 36 codes the synthesized

data and the coded data are fed to a multiplexer 37. The multiplexer 37 multiplexes an image from the image coder 36 with the voice data from the voice coder 31. A demultiplexer 41 demultiplexes the received multiplexed signal into the coded video image and the coded voice data, the coded video image is decoded by an image decoder 42 and the decoded image is displayed on a monitor 5, and the coded video data are decoded by a voice decoder 43 and the decoded voice data are outputted to a loudspeaker 6.



LEGAL STATUS

[Date of request for examination] 08.07.1998
[Date of sending the examiner's decision of rejection] 05.12.2000
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]
[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The picture transmission equipment characterized by to provide a synthetic means compound said data chosen from the speech-recognition means which carries out [voice / which was transmitted] speech recognition, an are-recording means are storing the data corresponding to the recognition result of said speech-recognition means, a selection means choose said data from said are-recording means, and said selection means, and said image which were transmitted, and to transmit the data corresponding to voice with an image. *1 pelet*

[Claim 2] Picture transmission equipment according to claim 1 characterized by accumulating the sign language image in said are recording means beforehand.

[Translation done.]

* NOTICES *

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] In a video conference system, this invention chooses data using speech recognition, and relates to the picture transmission equipment which compounds with an image and is transmitted.

[0002]

[Description of the Prior Art] Although a partner's image is generally seen during the board in many cases when using a video conference system, in the case of a content which is unclear, it is effective in employment of a board only with voice to use image information additionally. In the conventional video conference system, when image information, for example, a drawing etc., was sent as data and it was displayed as extra information, the data used beforehand had to be prepared.

[0003]

[Problem(s) to be Solved by the Invention] However, in the conventional video conference system, there was a problem that it was difficult to use timely the suitable image information according to expansion of a board. It is because multiplexing what was prepared beforehand, and the thing of the content which is not influenced by the information on the image transmitted or voice with an image and voice, and transmitting can only do the data which transmit the reason.

[0004] This invention was made in consideration of such a point, recognizes the voice which transmits by speech recognition at a television conference, and chooses the data corresponding to the voice from a database, and it aims at offering the picture transmission equipment which compounds with a transmitting image and is transmitted.

[0005]

[Means for Solving the Problem] A speech recognition means by which invention according to claim 1 carries out [voice / which was transmitted] speech recognition in a video conference system, An are recording means by which the data corresponding to the recognition result of said speech recognition means are stored, It is picture transmission equipment characterized by providing a synthetic means to compound the data chosen from a selection means to choose said data from said are recording means, and said selection means, and the transmitted image, and transmitting the data corresponding to voice with an image.

[0006]

[Embodiment of the Invention] Hereafter, 1 operation gestalt of this invention is explained with reference to a drawing. Drawing 1 is the block diagram showing the configuration of the image transmission equipment by 1 operation gestalt of this invention. In this drawing, 3 is the CODEC transmitting section and 4 is a CODEC receive section. The voice encoder which encodes the voice to which the sign 31 was supplied from the microphone 2 in the CODEC transmitting section 3, The speech recognition machine which recognizes the voice to which 32 was supplied from the microphone 2, the data judging machine which chooses from a database 34 the data corresponding to the voice 33 has been recognized to be with the speech recognition vessel 32, The synthetic vessel which compounds

the image to which 35 was supplied from the camera 1, and the data supplied from the database 34, The image encoder which encodes the image to which 36 was supplied from the synthetic vessel 35, and 37 are the coded image to which the coding voice supplied from the voice encoder 31 was supplied from the image encoder 36, and a multiplexing machine to multiplex. Moreover, in the CODEC receive section 4, they are the eliminator with which 41 divides the multiplexed signal of coding voice and a coded image into a coding image and coding voice, the image decryption machine which 42 decrypts the coded image from an eliminator 41, and is outputted to a monitor 5 through a line 425, and the voice decryption machine which 43 decrypts the coding voice from an eliminator 41, and is outputted to a loudspeaker 6 through a line 436.

[0007] Next, actuation of the picture transmission equipment by the above-mentioned configuration is explained. The image from a camera 1 is supplied to the synthetic vessel 35 through a line 135. It passes through the voice from a microphone 2 voice encoder 31 through a line 231, and it is supplied to the speech recognition machine 32 through a line 232. In the voice encoder 31, the supplied voice is encoded and the multiplexing machine 37 is supplied through a line 317. With the speech recognition vessel 32, the supplied voice is recognized and the data is supplied to the data judging machine 33 through a line 323. In the data judging machine 33, the data corresponding to the recognized voice are chosen from a database 34 through a line 343, and the judgment result is again supplied to a database 34 through a line 334. A database 34 supplies the data corresponding to the information supplied from the data judging machine 33 to the synthetic vessel 35 through a line 345. With the synthetic vessel 35, the image supplied from the camera 1 and the data supplied from the database 34 are compounded, and the image encoder 36 is supplied through a line 356. In the image encoder 36, the supplied image is encoded and the multiplexing machine 37 is supplied through a line 367. With the multiplexing vessel 37, the coding voice supplied from the coded image supplied from the image encoder 36 and the voice encoder 31 is multiplexed, and it outputs to a coupler.

[0008] In an eliminator 41, the multiplexed signal supplied through the coupler is divided into a coding image and coding voice, a coding image is supplied to the image decryption machine 42 through a line 412, and coding voice is supplied to the voice decryption machine 43 through a line 413. With the image decryption vessel 42, a coding image is decrypted and a monitor 5 is supplied through a line 425. The transmitted image is displayed in a monitor 5. With the voice decryption vessel 43, coding voice is decrypted and a loudspeaker 6 is supplied through a line 436. *monitor*

[0009] Drawing 2 is drawing showing the configuration of the video conference system constituted by two or more image transmission equipment. Transmission and reception of the data between two or more equipments are performed through networks, such as ISDN (Integrated Service Digital Network). The voice from the image and microphone 2 from a camera 1 by the side of Equipment A is supplied to the CODEC transmitting section 3. Such images and voice are supplied to a coupler 7 as data which the above processings were performed and were multiplexed in the CODEC transmitting section 3. The data supplied to the coupler 7 are supplied to coupler 7' by the side of Equipment B through ISDN8.

[0010] The data supplied to coupler 7' are supplied to CODEC receive section 4' by the side of Equipment B, as mentioned above, it separates into image data and voice data again, and are decrypted further, and are supplied to monitor 5' and loudspeaker 6' there, respectively. Although the above described the data flow from Equipment A side to Equipment B side, since the same is completely said of the data flow from Equipment B side to Equipment A side, the explanation about this is omitted.

[0011] When the image data of sign language are stored in the database by this invention and a data judging machine chooses from a database the image data of the sign language corresponding to the voice data transmitted from the speech recognition machine, it becomes possible to tell a person hard of hearing the conversation of those who do not know sign language. Moreover, when the data of foreign country language are stored in the database and a data judging machine chooses the translation word corresponding to the language to transmit, the conversation of those which speak different language is attained.

[0012]

[Effect of the Invention] Since according to this invention recognize the voice transmitted with a speech

recognition means, the data corresponding to the voice are chosen from an are recording means, it compounds with the transmitted image and it transmits as explained above, the effectiveness that the data corresponding to the voice of a transmitting side can be displayed on voice and coincidence at a receiving side is acquired.

[Translation done.]

* NOTICES *

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

TECHNICAL FIELD

[Field of the Invention] In a video conference system, this invention chooses data using speech recognition, and relates to the picture transmission equipment which compounds with an image and is transmitted.

[Translation done.]

*** NOTICES ***

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

PRIOR ART

[Description of the Prior Art] Although a partner's image is generally seen during the board in many cases when using a video conference system, in the case of a content which is unclear, it is effective in employment of a board only with voice to use image information additionally. In the conventional video conference system, when image information, for example, a drawing etc., was sent as data and it was displayed as extra information, the data used beforehand had to be prepared.

[Translation done.]

*** NOTICES ***

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

EFFECT OF THE INVENTION

[Effect of the Invention] Since according to this invention recognize the voice transmitted with a speech recognition means, the data corresponding to the voice are chosen from an are recording means, it compounds with the transmitted image and it transmits as explained above, the effectiveness that the data corresponding to the voice of a transmitting side can be displayed on voice and coincidence at a receiving side is acquired.

[Translation done.]

*** NOTICES ***

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

TECHNICAL PROBLEM

[Problem(s) to be Solved by the Invention] However, in the conventional video conference system, there was a problem that it was difficult to use timely the suitable image information according to expansion of a board. It is because multiplexing what was prepared beforehand, and the thing of the content which is not influenced by the information on the image transmitted or voice with an image and voice, and transmitting can only do the data which transmit the reason.

[0004] This invention was made in consideration of such a point, recognizes the voice which transmits by speech recognition at a television conference, and chooses the data corresponding to the voice from a database, and it aims at offering the picture transmission equipment which compounds with a transmitting image and is transmitted.

[Translation done.]

* NOTICES *

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram showing the configuration of the picture transmission equipment by 1 operation gestalt of this invention.

[Drawing 2] It is the block diagram which is twisted in 1 operation gestalt of this invention and in which showing an exchange of the data between two or more picture transmission equipments.

[Description of Notations]

- 1 and 1'. camera
- 2 and 2'. microphone
- 3 and 3'. CODEC transmitting section
- 31. Voice Encoder
- 32. Speech Recognition Machine
- 33. Data Judging Machine
- 34. Database
- 35. Synthetic Vessel
- 36. Image Encoder
- 37. Multiplexing Machine
- 4 and 4'. CODEC receive section
- 41. Eliminator
- 42. Image Decryption Machine
- 43. Voice Decryption Machine
- 5 and 5'. monitor
- 6 and 6'. loudspeaker
- 7 and 7'. coupler
- 8. ISDN

[Translation done.]

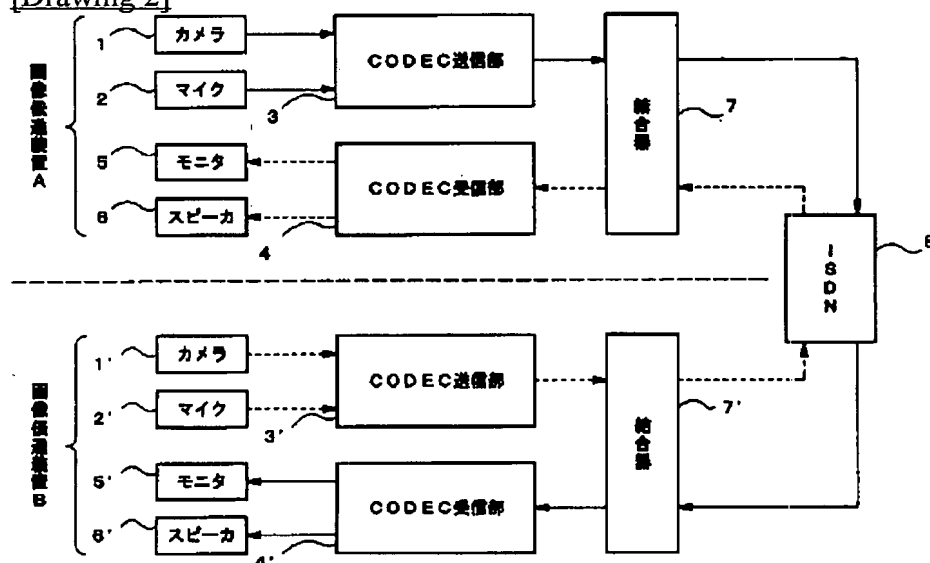
* NOTICES *

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DRAWINGS

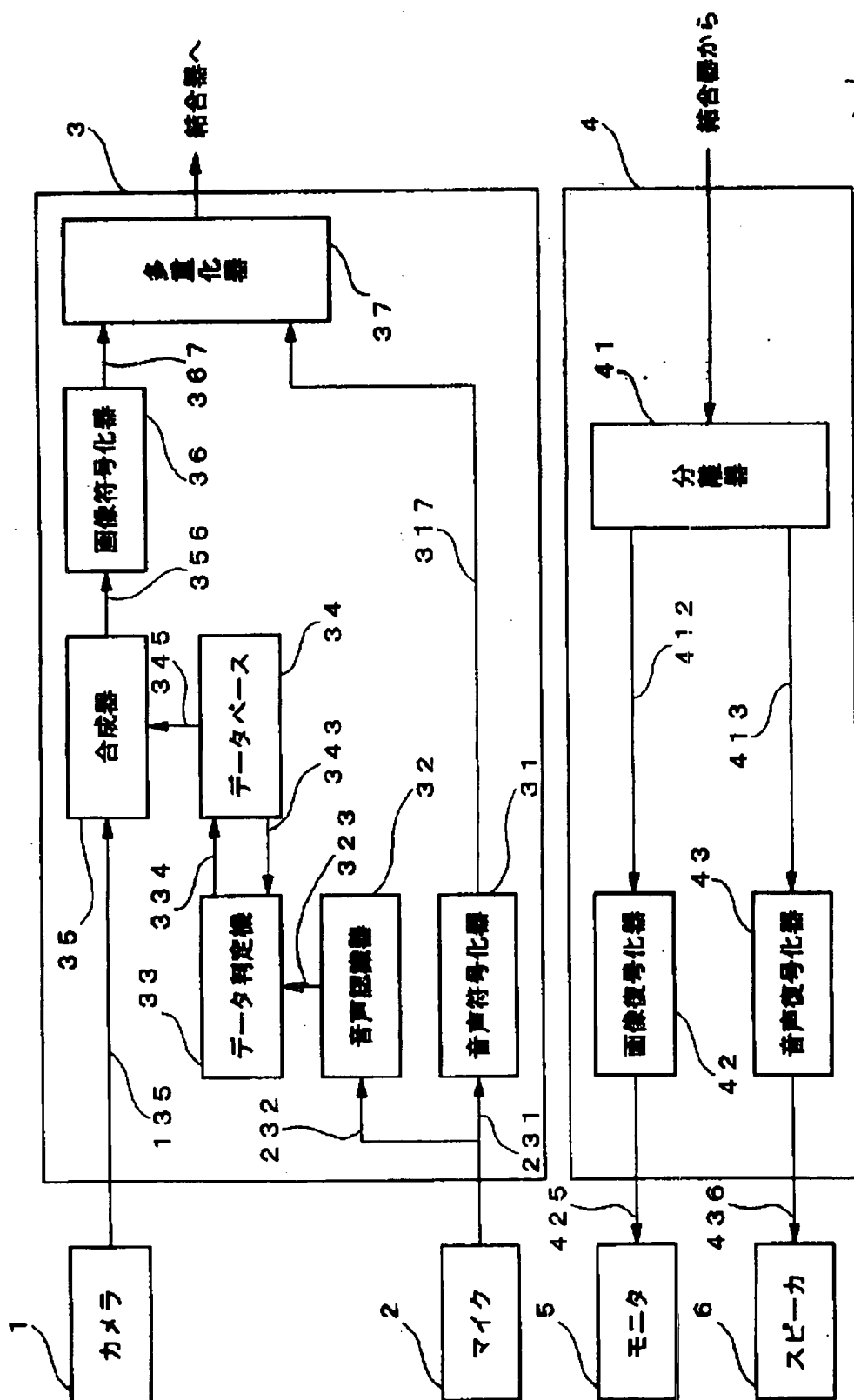
[Drawing 2]



1: Camera
2: mic
7: 7' / 7: Coupler

5': monitor
6': loudspeaker

[Drawing 1]



- 1: camera
- 2: microphone
- 3: Codec Transmitter
- 4: Codec receiver
- 5: monitor
- 6: loudspeaker

- 32: speech recognizer
- 34: database
- 36: image encoder
- 31: voice encoder
- 37: multiplexer
- 33: data judgment m/c

[Translation done.]